Supplementary Material for "Nonparametric Operator-Regularized Covariance Function Estimation for Functional Data"

Raymond K. W. Wong
Department of Statistics, Texas A&M University

Xiaoke Zhang
Department of Statistics, George Washington University

May 6, 2018

Abstract

This document provides the supplementary material to the article "Nonparametric Operator-Regularized Covariance Function Estimation for Functional Data" written by the same authors.

S1 Accelerated proximal gradient algorithm for Hilbert-Schmidt-norm regularization

The corresponding algorithm is the same as in Algorithm 1, except that we replace the proximal operator prox_ν in line 7 by

$$\text{prox}_\nu^{HS}(b) = \text{svec}\left[\underset{D \in \mathcal{S}_q^+}{\arg\min} \left\{ \frac{1}{2} \|D - B\|_F^2 + \nu \|D\|_F^2 \right\}\right]$$

The closed-form expression of such operator is given as follows. For any $\nu > 0$ and $b \in \mathbb{R}^{q(q+1)/2}$ with eigen-decomposition $\text{svec}^{-1}(b) = P\text{diag}(\tilde{b})P^\top$,

$$\text{prox}_\nu^{HS}(b) = \text{svec}(P\text{diag}(\tilde{c})P^\top),$$

where $\tilde{c} = (w_\nu(\tilde{b}_1), \ldots, w_\nu(\tilde{b}_q))^\top \in \mathbb{R}^q$ and $\tilde{b} = (\tilde{b}_1, \ldots, \tilde{b}_q)^\top \in \mathbb{R}^q$. Here $w_\nu(x) = (x/(1+2\nu))_+$ for any $x \in \mathbb{R}$.

S2 Additional simulation results

In this section, we first report full simulation results in Tables S1, S2 and S3. As mentioned in Section 4 in the article, the statistics for $\hat{C}^{+}_{\mathsf{SC}}$ are computed only based on successful runs, that is, those simulation runs where its corresponding package does not return an output due to computational errors are not counted, with the proportion of successful runs additionally shown in square brackets. In addition, the fpca package for computing $\hat{C}^{+}_{\mathsf{PP}}$ failed to provide an output in some simulation runs. We also noticed that although an output was obtained successfully, the estimator returned by the package could still be very unstable in some other simulation runs, which leads to a very large integrated squared error (ISE). Therefore, the results for $\hat{C}^{+}_{\mathsf{PP}}$ in Tables S1, S2 and S3 were calculated based on the remaining runs after the routine removal of the unsuccessful runs (i.e., with no output) and additional 5% runs (i.e., 15 runs) with the largest ISEs.

In addition, we performed another simulation study. The simulation settings are the same as those in the article except that the error variance $\sigma^2 = 0.1$ is higher, and only for $n = 50$ or 200. The corresponding results are given in Tables S4, S5 and S6, and we can reach similar conclusions to the article regarding the performance of these covariance estimators.

Table S1: AISE ($\times 10^3$) values with standard errors ($\times 10^3$) in parentheses for the ten covariance estimators, and average ranks with standard errors for those estimators with rank reduction. For $\hat{C}^+_{\mathsf{SC}}$, the percentage in a square bracket refers to the proportion of successful runs. For $\hat{C}^+_{\mathsf{PP}}$, the percentage in a square bracket refers to the proportion of successful runs minus 5%.

L	n	m		$\hat{C}^+_{\mathsf{trace}}$	$\hat{C}_{\mathsf{trace}}$	$\hat{C}^+_{\mathsf{HS}}$	$\hat{C}_{\mathsf{HS}}$	$\hat{C}_{\mathsf{CY}}$	$\hat{C}_{\mathsf{PACE}}$	$\hat{C}^+_{\mathsf{PACE,BIC}}$	$\hat{C}^+_{\mathsf{FACE}}$	$\hat{C}^+_{\mathsf{SC}}$	$\hat{C}^+_{\mathsf{PP}}$
2	50	10	AISE	9.25 (0.327)	11.55 (0.325)	9.00 (0.310)	12.32 (0.328)	10.26 (0.307)	12.55 (0.311)	11.94 (0.314)	9.37 (0.616)	10.64 (0.364)	4.38 (0.166) [88.7%]
			RANK	2.6 (0.031)	5.9 (0.150)	13.4 (0.049)	-	-	-	5.2 (0.047)	4.0 (0.036)	4.2 (0.039)	3.2 (0.036) [88.7%]
2	50	20	AISE	6.88 (0.290)	7.66 (0.266)	6.70 (0.260)	8.07 (0.263)	7.77 (0.263)	9.84 (0.255)	9.54 (0.253)	9.73 (0.967)	7.59 (0.296)	3.73 (0.164) [89.7%]
			RANK	2.6 (0.051)	7.5 (0.158)	13.0 (0.051)	-	-	-	4.5 (0.044)	3.7 (0.033)	3.9 (0.037)	3.4 (0.036) [89.7%]
2	100	10	AISE	5.26 (0.189)	6.69 (0.196)	5.13 (0.176)	7.08 (0.186)	6.43 (0.179)	8.76 (0.198)	8.37 (0.197)	6.01 (0.415)	5.94 (0.214)	2.39 (0.099) [93.7%]
			RANK	2.6 (0.032)	7.1 (0.116)	13.9 (0.047)	-	-	-	5.1 (0.046)	4.0 (0.032)	4.2 (0.038)	3.3 (0.039) [93.7%]
2	100	20	AISE	3.59 (0.135)	4.02 (0.137)	3.57 (0.134)	4.42 (0.129)	4.85 (0.133)	7.07 (0.165)	6.81 (0.161)	4.51 (0.340)	4.05 (0.141)	1.82 (0.083) [93.0%]
			RANK	2.6 (0.031)	7.8 (0.111)	13.6 (0.047)	-	-	-	4.5 (0.044)	3.9 (0.029)	3.9 (0.034)	3.2 (0.043) [93.0%]
2	200	10	AISE	2.85 (0.092)	3.55 (0.095)	2.81 (0.089)	4.08 (0.090)	4.30 (0.082)	6.69 (0.130)	6.37 (0.126)	3.40 (0.294)	3.18 (0.098)	1.06 (0.046) [93.7%]
			RANK	2.7 (0.033)	8.0 (0.121)	14.5 (0.043)	-	-	-	4.9 (0.045)	4.1 (0.033)	4.1 (0.037)	3.2 (0.045) [93.7%]
2	200	20	AISE	2.07 (0.078)	2.27 (0.078)	2.04 (0.077)	2.51 (0.079)	3.46 (0.079)	5.81 (0.116)	5.56 (0.112)	3.58 (0.393)	2.23 (0.080)	0.95 (0.045) [93.3%]
			RANK	2.7 (0.036)	8.1 (0.222)	14.3 (0.045)	-	-	-	4.4 (0.044)	4.0 (0.025)	3.9 (0.034)	3.0 (0.053) [93.3%]
4	50	10	AISE	17.04 (0.416)	20.15 (0.387)	15.94 (0.395)	21.09 (0.365)	17.86 (0.418)	16.43 (0.329)	15.56 (0.335)	14.69 (0.555)	16.09 (0.516) [99.7%]	> 1000 [88.0%]
			RANK	3.1 (0.055)	6.4 (0.269)	13.8 (0.047)	-	-	-	5.7 (0.049)	5.0 (0.036)	5.1 (0.039) [99.7%]	5.4 (0.049) [88.0%]
4	50	20	AISE	11.64 (0.338)	14.97 (0.311)	11.27 (0.317)	14.98 (0.285)	12.36 (0.316)	12.69 (0.295)	12.34 (0.294)	12.87 (0.869)	11.42 (0.360)	23.40 (2.095) [74.7%]
			RANK	3.8 (0.045)	13.2 (0.614)	13.5 (0.048)	-	-	-	5.2 (0.047)	5.2 (0.029)	5.0 (0.032)	5.5 (0.049) [74.7%]
4	100	10	AISE	9.44 (0.265)	13.56 (0.233)	8.89 (0.252)	13.95 (0.209)	11.37 (0.229)	11.44 (0.229)	10.92 (0.228)	9.25 (0.583)	8.77 (0.285)	> 1000 [84.7%]
			RANK	4.1 (0.046)	11.6 (0.558)	14.4 (0.042)	-	-	-	5.6 (0.051)	5.4 (0.032)	5.3 (0.032)	5.5 (0.045) [84.7%]
4	100	20	AISE	6.18 (0.167)	9.46 (0.171)	5.98 (0.162)	9.60 (0.168)	8.27 (0.163)	9.59 (0.182)	9.28 (0.178)	7.42 (0.588)	5.98 (0.166)	4.86 (0.777) [66.7%]
			RANK	4.4 (0.031)	26.0 (0.564)	14.3 (0.046)	-	-	-	5.1 (0.047)	5.4 (0.029)	5.1 (0.031)	5.5 (0.053) [66.7%]
4	200	10	AISE	4.94 (0.107)	9.01 (0.110)	4.74 (0.097)	9.08 (0.113)	7.25 (0.104)	9.01 (0.139)	8.61 (0.136)	6.50 (0.495)	4.64 (0.104)	5.50 (1.026) [84.0%]
			RANK	4.4 (0.032)	24.6 (0.624)	14.9 (0.046)	-	-	-	5.5 (0.047)	5.6 (0.030)	5.4 (0.032)	5.5 (0.046) [84.0%]
4	200	20	AISE	3.27 (0.081)	6.33 (0.081)	3.20 (0.080)	6.33 (0.081)	6.04 (0.076)	7.88 (0.121)	7.58 (0.116)	4.70 (0.295)	3.14 (0.081)	1.45 (0.064) [57.3%]
			RANK	4.5 (0.029)	31.9 (0.082)	15.0 (0.044)	-	-	-	5.0 (0.042)	5.7 (0.030)	5.1 (0.030)	5.4 (0.056) [57.3%]
10	50	10	AISE	19.99 (0.491)	23.81 (0.439)	18.45 (0.420)	24.59 (0.366)	21.11 (0.562)	18.86 (0.399)	17.74 (0.409)	18.70 (0.847)	20.83 (0.718) [99.7%]	> 1000 [93.7%]
			RANK	3.1 (0.054)	6.1 (0.259)	14.3 (0.045)	-	-	-	6.1 (0.048)	5.2 (0.041)	6.0 (0.036) [99.7%]	6.4 (0.068) [93.7%]
10	50	20	AISE	15.99 (0.398)	19.12 (0.367)	15.20 (0.370)	19.37 (0.354)	15.74 (0.371)	15.28 (0.308)	14.84 (0.308)	20.08 (1.754)	16.00 (0.445)	39.94 (9.631) [93.7%]
			RANK	3.7 (0.061)	9.7 (0.515)	14.5 (0.045)	-	-	-	6.1 (0.045)	5.6 (0.040)	6.7 (0.033)	6.8 (0.046) [93.7%]
10	100	10	AISE	12.72 (0.252)	17.30 (0.233)	12.16 (0.242)	17.83 (0.215)	14.62 (0.226)	13.96 (0.246)	13.27 (0.246)	13.01 (0.728)	12.21 (0.279) [99.0%]	> 1000 [92.3%]
			RANK	3.9 (0.056)	9.4 (0.487)	15.3 (0.048)	-	-	-	6.2 (0.047)	5.8 (0.039)	6.5 (0.037) [99.0%]	6.8 (0.047) [92.3%]
10	100	20	AISE	9.06 (0.179)	12.61 (0.175)	8.58 (0.175)	12.74 (0.175)	10.23 (0.156)	11.31 (0.185)	10.96 (0.182)	12.94 (1.050)	8.49 (0.188)	6.13 (0.463) [91.0%]
			RANK	4.6 (0.052)	21.8 (0.690)	15.6 (0.044)	-	-	-	6.3 (0.044)	6.4 (0.038)	7.2 (0.034)	7.0 (0.021) [91.0%]
10	200	10	AISE	8.08 (0.158)	12.34 (0.155)	7.54 (0.144)	12.52 (0.153)	9.99 (0.149)	10.89 (0.163)	10.41 (0.160)	10.35 (0.694)	7.14 (0.166) [96.0%]	149.12 (75.357) [91.3%]
			RANK	4.7 (0.052)	18.8 (0.693)	15.9 (0.048)	-	-	-	6.4 (0.045)	6.4 (0.036)	6.9 (0.035) [96.0%]	6.9 (0.027) [91.3%]
10	200	20	AISE	5.42 (0.099)	8.81 (0.098)	5.09 (0.087)	8.81 (0.096)	7.85 (0.088)	9.55 (0.128)	9.23 (0.124)	10.52 (0.611)	4.57 (0.096)	2.70 (0.058) [93.7%]
			RANK	5.8 (0.068)	31.2 (0.255)	16.5 (0.044)	-	-	-	6.5 (0.040)	7.2 (0.031)	7.7 (0.030)	7.0 (0.004) [93.7%]

Table S2: Bias ($\times 10^2$) and MSE ($\times 10^4$) values with their standard errors (multiplied by 10^2 and 10^4 respectively) in parentheses for the principal eigenvalue ζ_1, and AISE ($\times 10^2$) values with standard errors ($\times 10^2$) for the principal eigenfunction ϕ_1. The statistics for $\hat{C}^+_{\mathsf{SC}}$ and $\hat{C}^+_{\mathsf{PP}}$ are computed similarly as described in Table S1.

L	n	m		$\hat{C}^+_{\mathsf{trace}}$	$\hat{C}^+_{\mathsf{HS}}$	$\hat{C}^+_{\mathsf{PACE,BIC}}$	$\hat{C}^+_{\mathsf{FACE}}$	$\hat{C}^+_{\mathsf{SC}}$	$\hat{C}^+_{\mathsf{PP}}$
2	50	10	ζ_1 (BIAS)	-1.159 (0.30)	-1.570 (0.29)	-4.823 (0.26)	-0.799 (0.35)	-0.507 (0.30)	-0.973 (0.27) [88.7%]
			ζ_1 (MSE)	27.62 (2.04)	27.70 (1.99)	44.06 (2.63)	37.67 (4.59)	27.24 (2.06)	19.85 (1.38) [88.7%]
			ϕ_1 (AISE)	6.71 (0.416)	6.58 (0.404)	7.06 (0.465)	6.73 (0.589)	7.70 (0.437)	3.18 (0.257) [88.7%]
2	50	20	ζ_1 (BIAS)	-1.166 (0.29)	-1.464 (0.29)	-4.971 (0.25)	-0.188 (0.41)	-0.675 (0.30)	-1.141 (0.26) [89.7%]
			ζ_1 (MSE)	27.16 (2.32)	26.71 (2.23)	43.13 (2.43)	49.96 (6.64)	26.89 (2.34)	19.01 (1.39) [89.7%]
			ϕ_1 (AISE)	5.53 (0.410)	5.34 (0.378)	4.85 (0.263)	5.89 (0.509)	6.09 (0.448)	3.23 (0.276) [89.7%]
2	100	10	ζ_1 (BIAS)	-0.912 (0.23)	-1.098 (0.23)	-4.647 (0.21)	-0.409 (0.28)	-0.607 (0.24)	-0.760 (0.20) [93.7%]
			ζ_1 (MSE)	17.15 (1.33)	16.83 (1.27)	34.55 (1.87)	24.22 (2.92)	17.02 (1.36)	12.04 (0.85) [93.7%]
			ϕ_1 (AISE)	3.44 (0.182)	3.38 (0.177)	3.93 (0.163)	3.42 (0.224)	3.77 (0.194)	1.83 (0.141) [93.7%]
2	100	20	ζ_1 (BIAS)	-0.370 (0.21)	-0.487 (0.21)	-4.369 (0.17)	0.471 (0.27)	-0.064 (0.21)	-0.307 (0.18) [93.0%]
			ζ_1 (MSE)	12.88 (0.99)	12.84 (0.98)	28.01 (1.54)	21.26 (2.66)	12.97 (1.02)	8.80 (0.67) [93.0%]
			ϕ_1 (AISE)	2.50 (0.166)	2.48 (0.158)	2.92 (0.122)	2.50 (0.192)	2.71 (0.174)	1.43 (0.123) [93.0%]
2	200	10	ζ_1 (BIAS)	-0.465 (0.16)	-0.567 (0.16)	-4.423 (0.14)	0.052 (0.21)	-0.295 (0.16)	-0.323 (0.13) [93.7%]
			ζ_1 (MSE)	7.90 (0.67)	7.91 (0.66)	25.14 (1.25)	12.72 (2.05)	7.88 (0.67)	5.08 (0.36) [93.7%]
			ϕ_1 (AISE)	1.84 (0.079)	1.82 (0.078)	2.56 (0.083)	1.66 (0.129)	1.90 (0.082)	0.79 (0.063) [93.7%]
2	200	20	ζ_1 (BIAS)	-0.120 (0.15)	-0.175 (0.15)	-4.252 (0.13)	1.034 (0.24)	0.061 (0.16)	0.111 (0.13) [93.3%]
			ζ_1 (MSE)	7.17 (0.56)	7.19 (0.56)	23.19 (1.11)	18.61 (3.27)	7.26 (0.57)	5.00 (0.39) [93.3%]
			ϕ_1 (AISE)	1.32 (0.066)	1.30 (0.065)	2.04 (0.063)	1.55 (0.137)	1.35 (0.067)	0.71 (0.056) [93.3%]
4	50	10	ζ_1 (BIAS)	-0.744 (0.33)	-1.131 (0.33)	-4.374 (0.30)	-0.308 (0.36)	0.227 (0.34) [99.7%]	< -1000 or > 1000 [88.0%]
			ζ_1 (MSE)	33.81 (2.82)	34.05 (2.88)	45.22 (2.62)	38.37 (3.91)	34.97 (3.18) [99.7%]	> 1000 [88.0%]
			ϕ_1 (AISE)	10.23 (0.784)	9.96 (0.787)	9.50 (0.674)	9.96 (0.892)	11.35 (0.873) [99.7%]	72.71 (5.207) [88.0%]
4	50	20	ζ_1 (BIAS)	-0.485 (0.29)	-0.761 (0.29)	-4.865 (0.25)	0.864 (0.39)	-0.021 (0.30)	-5.835 (0.67) [74.7%]
			ζ_1 (MSE)	25.90 (2.16)	25.51 (2.06)	43.03 (2.50)	45.34 (6.26)	27.52 (2.39)	133.89 (14.22) [74.7%]
			ϕ_1 (AISE)	7.66 (0.481)	7.41 (0.477)	6.34 (0.317)	7.84 (0.637)	8.39 (0.484)	45.14 (4.973) [74.7%]
4	100	10	ζ_1 (BIAS)	-0.356 (0.23)	-0.507 (0.23)	-4.624 (0.20)	0.639 (0.30)	-0.226 (0.24)	31.491 (26.35) [84.7%]
			ζ_1 (MSE)	16.04 (1.51)	15.74 (1.47)	33.87 (2.01)	26.52 (4.95)	16.66 (1.67)	> 1000 [84.7%]
			ϕ_1 (AISE)	5.06 (0.233)	4.76 (0.221)	4.92 (0.210)	4.28 (0.263)	5.18 (0.240)	52.12 (5.239) [84.7%]
4	100	20	ζ_1 (BIAS)	-0.103 (0.21)	-0.132 (0.21)	-4.480 (0.18)	1.143 (0.27)	-0.038 (0.21)	-0.764 (0.28) [66.7%]
			ζ_1 (MSE)	13.31 (1.16)	13.29 (1.15)	30.07 (1.67)	22.44 (4.24)	13.33 (1.15)	16.48 (3.40) [66.7%]
			ϕ_1 (AISE)	3.95 (0.195)	3.83 (0.192)	3.85 (0.155)	4.91 (0.627)	4.03 (0.197)	5.88 (1.543) [66.7%]
4	200	10	ζ_1 (BIAS)	-0.347 (0.16)	-0.370 (0.16)	-4.665 (0.14)	0.896 (0.24)	-0.380 (0.16)	-1.398 (0.32) [84.0%]
			ζ_1 (MSE)	7.87 (0.61)	7.80 (0.61)	27.61 (1.30)	17.47 (3.64)	7.80 (0.61)	27.14 (6.63) [84.0%]
			ϕ_1 (AISE)	2.71 (0.109)	2.63 (0.104)	3.25 (0.110)	3.44 (0.671)	2.67 (0.108)	8.63 (2.118) [84.0%]
4	200	20	ζ_1 (BIAS)	-0.142 (0.15)	-0.157 (0.15)	-4.482 (0.13)	1.224 (0.21)	-0.125 (0.15)	-0.522 (0.16) [57.3%]
			ζ_1 (MSE)	7.01 (0.59)	6.98 (0.58)	25.28 (1.15)	15.01 (2.34)	6.99 (0.58)	4.85 (0.46) [57.3%]
			ϕ_1 (AISE)	1.90 (0.086)	1.84 (0.085)	2.45 (0.073)	2.24 (0.177)	1.88 (0.087)	1.10 (0.097) [57.3%]
10	50	10	ζ_1 (BIAS)	-0.993 (0.34)	-1.584 (0.33)	-4.660 (0.31)	-0.419 (0.40)	0.419 (0.36) [99.7%]	< -1000 or > 1000 [93.7%]
			ζ_1 (MSE)	36.49 (2.95)	35.52 (3.15)	49.75 (3.05)	49.11 (6.06)	37.98 (4.08) [99.7%]	> 1000 [93.7%]
			ϕ_1 (AISE)	10.74 (0.685)	10.44 (0.736)	9.97 (0.688)	10.56 (0.850)	13.19 (0.788) [99.7%]	123.18 (4.802) [93.7%]
10	50	20	ζ_1 (BIAS)	-0.176 (0.32)	-0.509 (0.32)	-4.430 (0.28)	1.506 (0.52)	0.635 (0.33)	-2.004 (0.64) [93.7%]
			ζ_1 (MSE)	31.51 (2.48)	31.50 (2.44)	42.67 (2.61)	84.57 (14.41)	32.39 (2.66)	117.44 (33.06) [93.7%]
			ϕ_1 (AISE)	9.53 (0.740)	9.41 (0.765)	7.85 (0.520)	8.97 (0.726)	11.58 (0.874)	36.47 (3.890) [93.7%]
10	100	10	ζ_1 (BIAS)	-0.329 (0.25)	-0.477 (0.25)	-4.639 (0.22)	0.647 (0.30)	0.133 (0.25) [99.0%]	< -1000 or > 1000 [92.3%]
			ζ_1 (MSE)	18.25 (1.39)	18.64 (1.46)	35.72 (2.03)	28.23 (4.61)	18.89 (1.51) [99.0%]	> 1000 [92.3%]
			ϕ_1 (AISE)	6.22 (0.332)	6.10 (0.325)	6.13 (0.311)	7.55 (0.777)	7.02 (0.342) [99.0%]	83.44 (5.300) [92.3%]
10	100	20	ζ_1 (BIAS)	0.258 (0.22)	0.199 (0.22)	-4.237 (0.19)	2.438 (0.38)	0.447 (0.22)	0.110 (0.23) [91.0%]
			ζ_1 (MSE)	14.32 (1.13)	14.03 (1.07)	28.87 (1.73)	48.81 (8.59)	14.44 (1.14)	14.88 (2.22) [91.0%]
			ϕ_1 (AISE)	4.39 (0.213)	4.23 (0.209)	4.11 (0.166)	5.19 (0.373)	4.82 (0.224)	4.64 (0.822) [91.0%]
10	200	10	ζ_1 (BIAS)	0.061 (0.19)	0.036 (0.18)	-4.396 (0.16)	1.514 (0.26)	0.065 (0.19) [96.0%]	0.171 (1.61) [91.3%]
			ζ_1 (MSE)	10.26 (0.99)	10.16 (0.97)	26.85 (1.44)	22.57 (3.79)	10.10 (0.98) [96.0%]	710.65 (466.54) [91.3%]
			ϕ_1 (AISE)	3.27 (0.136)	3.14 (0.131)	3.48 (0.148)	3.66 (0.355)	3.43 (0.143) [96.0%]	37.85 (4.258) [91.3%]
10	200	20	ζ_1 (BIAS)	0.110 (0.17)	0.033 (0.17)	-4.397 (0.14)	2.791 (0.26)	0.126 (0.17)	0.029 (0.14) [93.7%]
			ζ_1 (MSE)	8.20 (0.64)	8.23 (0.64)	25.01 (1.21)	28.75 (3.72)	8.29 (0.65)	5.71 (0.41) [93.7%]
			ϕ_1 (AISE)	2.49 (0.104)	2.21 (0.096)	2.74 (0.089)	6.25 (0.954)	2.42 (0.101)	1.50 (0.070) [93.7%]

Table S3: Similar to Table S2, but for the second eigenvalue ζ_2 and second eigenfunction ϕ_2.

L	n	m		$\hat{C}^+_{\text{trace}}$	$\hat{C}^+_{\text{HS}}$	$\hat{C}^+_{\text{PACE,BIC}}$	$\hat{C}^+_{\text{FACE}}$	$\hat{C}^+_{\text{SC}}$	$\hat{C}^+_{\text{PP}}$
2	50	10	ζ_2 (BIAS)	-1.8531 (0.145)	-1.8349 (0.139)	-3.6440 (0.116)	-0.8609 (0.156)	-0.8940 (0.141)	-1.0369 (0.125) [88.7%]
			ζ_2 (MSE)	9.68 (0.620)	9.16 (0.573)	17.30 (0.801)	7.98 (0.621)	6.77 (0.487)	5.22 (0.390) [88.7%]
			ϕ_2 (AISE)	9.64 (0.508)	9.06 (0.463)	11.90 (0.639)	9.39 (0.585)	10.90 (0.481)	3.91 (0.274) [88.7%]
2	50	20	ζ_2 (BIAS)	-1.3181 (0.126)	-1.3575 (0.122)	-3.6482 (0.095)	-0.4270 (0.170)	-0.6541 (0.127)	-0.7436 (0.122) [89.7%]
			ζ_2 (MSE)	6.47 (0.474)	6.26 (0.433)	16.01 (0.612)	8.83 (1.495)	5.26 (0.446)	4.56 (0.375) [89.7%]
			ϕ_2 (AISE)	6.65 (0.442)	6.38 (0.401)	6.54 (0.310)	7.38 (0.511)	7.73 (0.465)	3.51 (0.285) [89.7%]
2	100	10	ζ_2 (BIAS)	-0.8309 (0.106)	-0.8061 (0.100)	-3.1957 (0.082)	-0.0304 (0.120)	-0.2241 (0.102)	-0.4162 (0.090) [93.7%]
			ζ_2 (MSE)	4.04 (0.321)	3.66 (0.283)	12.24 (0.521)	4.31 (0.513)	3.16 (0.270)	2.46 (0.223) [93.7%]
			ϕ_2 (AISE)	4.39 (0.229)	4.18 (0.209)	5.23 (0.220)	5.03 (0.254)	5.06 (0.231)	2.12 (0.144) [93.7%]
2	100	20	ζ_2 (BIAS)	-0.8478 (0.092)	-0.8211 (0.090)	-3.4120 (0.070)	-0.1509 (0.106)	-0.4107 (0.093)	-0.5311 (0.089) [93.0%]
			ζ_2 (MSE)	3.27 (0.235)	3.11 (0.226)	13.12 (0.441)	3.41 (0.404)	2.76 (0.237)	2.49 (0.198) [93.0%]
			ϕ_2 (AISE)	2.98 (0.179)	2.89 (0.169)	3.47 (0.147)	3.67 (0.201)	3.51 (0.181)	1.57 (0.130) [93.0%]
2	200	10	ζ_2 (BIAS)	-0.5193 (0.076)	-0.4813 (0.074)	-3.1322 (0.060)	0.1884 (0.092)	-0.1423 (0.076)	-0.2778 (0.066) [93.7%]
			ζ_2 (MSE)	2.00 (0.145)	1.89 (0.137)	10.90 (0.371)	2.58 (0.417)	1.74 (0.127)	1.30 (0.098) [93.7%]
			ϕ_2 (AISE)	2.40 (0.101)	2.31 (0.095)	3.17 (0.116)	3.09 (0.167)	2.75 (0.098)	0.89 (0.064) [93.7%]
2	200	20	ζ_2 (BIAS)	-0.3549 (0.072)	-0.3120 (0.071)	-3.1451 (0.053)	0.3658 (0.094)	-0.0824 (0.072)	-0.1817 (0.066) [93.3%]
			ζ_2 (MSE)	1.66 (0.122)	1.59 (0.117)	10.74 (0.328)	2.75 (0.420)	1.55 (0.124)	1.23 (0.100) [93.3%]
			ϕ_2 (AISE)	1.50 (0.072)	1.44 (0.069)	1.69 (0.065)	2.33 (0.139)	1.71 (0.070)	0.76 (0.058) [93.3%]
4	50	10	ζ_2 (BIAS)	-1.4561 (0.152)	-1.3670 (0.154)	-3.1917 (0.117)	-0.2336 (0.181)	-0.0567 (0.147) [99.7%]	59.5874 (43.186) [88.0%]
			ζ_2 (MSE)	8.98 (0.627)	8.95 (0.629)	14.29 (0.697)	9.84 (1.100)	6.48 (0.498) [99.7%]	> 1000 [88.0%]
			ϕ_2 (AISE)	23.31 (1.542)	19.56 (1.398)	26.02 (1.734)	17.47 (1.250)	29.40 (1.890) [99.7%]	84.32 (5.094) [88.0%]
4	50	20	ζ_2 (BIAS)	-0.5721 (0.140)	-0.6102 (0.139)	-3.2647 (0.101)	0.5130 (0.186)	0.1033 (0.137)	-2.4269 (0.300) [74.7%]
			ζ_2 (MSE)	6.16 (0.474)	6.19 (0.462)	13.72 (0.638)	10.57 (1.831)	5.65 (0.465)	26.01 (2.856) [74.7%]
			ϕ_2 (AISE)	14.95 (0.796)	13.45 (0.726)	12.99 (0.786)	13.30 (0.884)	20.66 (1.349)	56.66 (5.042) [74.7%]
4	100	10	ζ_2 (BIAS)	-0.3774 (0.123)	-0.3662 (0.120)	-3.1653 (0.090)	0.5736 (0.141)	0.0203 (0.116)	-2.4787 (0.293) [84.7%]
			ζ_2 (MSE)	4.69 (0.421)	4.45 (0.379)	12.43 (0.541)	6.31 (0.713)	4.00 (0.355)	27.80 (3.836) [84.7%]
			ϕ_2 (AISE)	11.83 (0.906)	10.19 (0.807)	10.42 (0.558)	9.27 (0.701)	14.21 (0.985)	53.66 (4.862) [84.7%]
4	100	20	ζ_2 (BIAS)	-0.2117 (0.104)	-0.1758 (0.104)	-3.2562 (0.080)	0.9431 (0.152)	-0.0274 (0.104)	-0.3523 (0.146) [66.7%]
			ζ_2 (MSE)	3.27 (0.240)	3.24 (0.242)	12.52 (0.509)	7.77 (1.520)	3.22 (0.248)	4.34 (0.983) [66.7%]
			ϕ_2 (AISE)	9.09 (0.527)	8.27 (0.455)	6.61 (0.375)	9.32 (0.678)	11.13 (0.639)	11.42 (1.896) [66.7%]
4	200	10	ζ_2 (BIAS)	-0.0793 (0.084)	-0.0298 (0.083)	-3.1236 (0.067)	1.2165 (0.126)	0.0178 (0.084)	-0.6063 (0.145) [84.0%]
			ζ_2 (MSE)	2.10 (0.177)	2.05 (0.177)	11.09 (0.393)	6.24 (0.882)	2.11 (0.182)	5.61 (1.327) [84.0%]
			ϕ_2 (AISE)	6.75 (0.476)	6.28 (0.368)	5.13 (0.205)	7.95 (0.696)	7.64 (0.498)	13.03 (2.431) [84.0%]
4	200	20	ζ_2 (BIAS)	-0.0383 (0.074)	-0.0142 (0.074)	-3.1518 (0.057)	1.1588 (0.103)	0.0324 (0.074)	-0.1192 (0.088) [57.3%]
			ζ_2 (MSE)	1.64 (0.130)	1.63 (0.130)	10.89 (0.351)	4.53 (0.552)	1.63 (0.133)	1.34 (0.132) [57.3%]
			ϕ_2 (AISE)	4.52 (0.230)	4.19 (0.199)	2.91 (0.128)	5.92 (0.288)	5.10 (0.260)	2.58 (0.195) [57.3%]
10	50	10	ζ_2 (BIAS)	-1.3218 (0.175)	-1.2629 (0.165)	-2.9539 (0.126)	-0.1583 (0.185)	0.5085 (0.163) [99.7%]	< −1000 or > 1000 [93.7%]
			ζ_2 (MSE)	10.93 (0.833)	9.75 (0.697)	13.46 (0.710)	10.30 (0.955)	8.13 (0.714) [99.7%]	> 1000 [93.7%]
			ϕ_2 (AISE)	25.07 (1.312)	20.71 (1.063)	25.05 (1.449)	18.91 (1.148)	36.40 (2.043) [99.7%]	130.32 (4.126) [93.7%]
10	50	20	ζ_2 (BIAS)	-0.9957 (0.147)	-1.0116 (0.145)	-3.4308 (0.106)	0.1638 (0.203)	0.2121 (0.137)	-0.9921 (0.214) [93.7%]
			ζ_2 (MSE)	7.43 (0.637)	7.31 (0.610)	15.14 (0.713)	12.36 (2.134)	5.66 (0.568)	13.78 (1.760) [93.7%]
			ϕ_2 (AISE)	21.68 (1.521)	18.67 (1.285)	18.49 (1.272)	15.91 (0.995)	33.72 (2.049)	52.18 (3.751) [93.7%]
10	100	10	ζ_2 (BIAS)	-0.2501 (0.121)	-0.2139 (0.118)	-2.9185 (0.093)	0.9163 (0.168)	0.5290 (0.115) [99.0%]	< −1000 or > 1000 [92.3%]
			ζ_2 (MSE)	4.46 (0.396)	4.24 (0.347)	11.11 (0.537)	9.27 (1.647)	4.18 (0.374) [99.0%]	> 1000 [92.3%]
			ϕ_2 (AISE)	15.60 (0.939)	14.06 (0.727)	14.45 (0.985)	13.56 (0.915)	22.13 (1.204) [99.0%]	91.96 (5.026) [92.3%]
10	100	20	ζ_2 (BIAS)	-0.1887 (0.108)	-0.1510 (0.106)	-3.2046 (0.078)	1.0801 (0.150)	0.1797 (0.106)	0.0071 (0.105) [91.0%]
			ζ_2 (MSE)	3.51 (0.323)	3.40 (0.313)	12.07 (0.474)	7.87 (0.998)	3.38 (0.339)	3.00 (0.403) [91.0%]
			ϕ_2 (AISE)	10.31 (0.495)	9.54 (0.467)	8.07 (0.448)	11.79 (0.695)	15.47 (0.901)	10.93 (1.107) [91.0%]
10	200	10	ζ_2 (BIAS)	0.0104 (0.089)	0.0705 (0.088)	-3.0623 (0.072)	1.6463 (0.176)	0.2391 (0.091) [96.0%]	-0.1692 (0.581) [91.3%]
			ζ_2 (MSE)	2.37 (0.202)	2.30 (0.196)	10.91 (0.428)	11.98 (2.216)	2.41 (0.218) [96.0%]	92.21 (69.849) [91.3%]
			ϕ_2 (AISE)	7.76 (0.349)	7.12 (0.302)	7.31 (0.361)	9.60 (0.708)	10.15 (0.438) [96.0%]	42.17 (4.377) [91.3%]
10	200	20	ζ_2 (BIAS)	0.0769 (0.069)	0.0432 (0.069)	-3.0610 (0.056)	2.2965 (0.156)	0.1473 (0.069)	0.0704 (0.064) [93.7%]
			ζ_2 (MSE)	1.44 (0.125)	1.43 (0.122)	10.30 (0.331)	12.60 (1.368)	1.45 (0.125)	1.15 (0.102) [93.7%]
			ϕ_2 (AISE)	6.25 (0.317)	5.22 (0.226)	4.00 (0.186)	12.36 (1.018)	7.13 (0.351)	4.21 (0.180) [93.7%]

Table S4: Similar to Table S1, but for the high error variance ($\sigma^2 = 0.1$) and $n = 50$ or 200.

L	n	m		$\hat{C}^{+}_{\mathsf{trace}}$	$\hat{C}_{\mathsf{trace}}$	$\hat{C}^{+}_{\mathsf{HS}}$	$\hat{C}_{\mathsf{HS}}$	$\hat{C}_{\mathsf{CY}}$	$\hat{C}_{\mathsf{PACE}}$	$\hat{C}^{+}_{\mathsf{PACE,BIC}}$	$\hat{C}^{+}_{\mathsf{FACE}}$	$\hat{C}^{+}_{\mathsf{SC}}$	$\hat{C}^{+}_{\mathsf{PP}}$
2	50	10	AISE	10.39 (0.351)	13.11 (0.350)	10.29 (0.343)	14.07 (0.361)	11.36 (0.344)	12.98 (0.298)	12.25 (0.298)	10.28 (0.595)	11.92 (0.415)	6.57 (0.211) [90.3%]
			rank	2.5 (0.033)	5.5 (0.133)	13.5 (0.047)	-	-	-	5.2 (0.045)	3.9 (0.037)	4.2 (0.041)	3.3 (0.053) [90.3%]
2	50	20	AISE	7.63 (0.293)	8.60 (0.258)	7.25 (0.245)	8.92 (0.241)	8.30 (0.248)	10.13 (0.269)	9.82 (0.268)	9.07 (1.116)	8.41 (0.292)	5.21 (0.211) [89.3%]
			rank	2.6 (0.040)	6.9 (0.144)	13.1 (0.051)	-	-	-	4.6 (0.045)	3.8 (0.036)	3.9 (0.038)	3.5 (0.058) [89.3%]
2	200	10	AISE	3.35 (0.109)	4.32 (0.113)	3.32 (0.103)	4.89 (0.109)	4.83 (0.104)	7.17 (0.139)	6.81 (0.136)	4.76 (0.493)	3.70 (0.114)	1.85 (0.060) [94.0%]
			rank	2.7 (0.034)	8.3 (0.173)	14.6 (0.047)	-	-	-	5.0 (0.045)	4.2 (0.031)	4.2 (0.038)	3.2 (0.063) [94.0%]
2	200	20	AISE	2.26 (0.079)	2.51 (0.080)	2.23 (0.078)	2.80 (0.080)	3.61 (0.078)	6.27 (0.133)	6.00 (0.128)	2.66 (0.170)	2.40 (0.078)	1.34 (0.048) [90.0%]
			rank	2.6 (0.031)	8.1 (0.199)	14.5 (0.040)	-	-	-	4.4 (0.039)	4.0 (0.026)	4.0 (0.032)	3.6 (0.074) [90.0%]
4	50	10	AISE	17.78 (0.467)	20.96 (0.369)	16.79 (0.417)	22.15 (0.373)	18.77 (0.472)	16.73 (0.327)	15.69 (0.327)	14.81 (0.453)	16.99 (0.582)	13.06 (0.449) [94.0%]
			rank	3.0 (0.048)	5.6 (0.151)	13.8 (0.043)	-	-	-	5.8 (0.048)	4.9 (0.037)	5.1 (0.035)	4.6 (0.054) [94.0%]
4	50	20	AISE	12.71 (0.385)	15.98 (0.310)	12.12 (0.335)	16.22 (0.306)	13.19 (0.318)	12.75 (0.281)	12.35 (0.280)	16.35 (2.703)	12.10 (0.380)	7.50 (0.225) [92.0%]
			rank	3.8 (0.049)	11.7 (0.578)	13.6 (0.045)	-	-	-	5.2 (0.048)	5.2 (0.029)	5.0 (0.035)	5.0 (0.054) [92.0%]
4	200	10	AISE	5.74 (0.124)	10.03 (0.127)	5.56 (0.124)	10.24 (0.133)	8.30 (0.116)	9.29 (0.151)	8.84 (0.149)	6.89 (0.434)	5.41 (0.134)	2.60 (0.063) [91.3%]
			rank	4.4 (0.034)	20.3 (0.690)	15.0 (0.045)	-	-	-	5.5 (0.048)	5.6 (0.033)	5.3 (0.034)	4.7 (0.046) [91.3%]
4	200	20	AISE	3.58 (0.098)	6.74 (0.100)	3.48 (0.095)	6.73 (0.099)	6.23 (0.098)	7.97 (0.121)	7.67 (0.117)	4.43 (0.291)	3.41 (0.096)	1.78 (0.049) [88.0%]
			rank	4.6 (0.036)	31.5 (0.180)	15.1 (0.041)	-	-	-	5.1 (0.046)	5.7 (0.029)	5.2 (0.031)	4.8 (0.058) [88.0%]
10	50	10	AISE	21.40 (0.497)	25.46 (0.488)	20.51 (0.482)	27.13 (0.490)	23.27 (0.667)	20.48 (0.375)	18.96 (0.380)	19.46 (0.929)	22.79 (0.704)	> 1000 [87.3%]
			rank	3.0 (0.053)	5.6 (0.204)	14.2 (0.044)	-	-	-	6.1 (0.051)	5.1 (0.041)	5.8 (0.038)	5.3 (0.085) [87.3%]
10	50	20	AISE	15.86 (0.353)	18.94 (0.318)	14.85 (0.332)	19.45 (0.310)	15.79 (0.339)	15.30 (0.296)	14.78 (0.297)	20.90 (2.835)	16.15 (0.432)	19.97 (1.171) [94.0%]
			rank	3.7 (0.063)	9.0 (0.462)	14.4 (0.043)	-	-	-	6.0 (0.046)	5.6 (0.040)	6.5 (0.034)	6.5 (0.070) [94.0%]
10	200	10	AISE	8.41 (0.163)	13.09 (0.160)	7.85 (0.146)	13.28 (0.144)	10.85 (0.137)	11.28 (0.177)	10.77 (0.174)	11.19 (0.767)	7.37 (0.153)	8.47 (0.780) [90.3%]
			rank	4.6 (0.051)	14.8 (0.645)	15.9 (0.046)	-	-	-	6.4 (0.047)	6.3 (0.036)	6.9 (0.034)	6.7 (0.041) [90.3%]
10	200	20	AISE	5.66 (0.114)	9.09 (0.111)	5.29 (0.110)	9.08 (0.109)	7.88 (0.102)	9.62 (0.131)	9.29 (0.128)	10.91 (0.653)	4.80 (0.112)	3.46 (0.067) [93.3%]
			rank	5.6 (0.069)	31.0 (0.279)	16.4 (0.041)	-	-	-	6.4 (0.041)	7.1 (0.031)	7.6 (0.030)	7.0 (0.004) [93.3%]

Table S5: Similar to Table S2, but for the high error variance ($\sigma^2 = 0.1$) and $n = 50$ or 200.

L	n	m		$\hat{C}^+_{\mathsf{trace}}$	$\hat{C}^+_{\mathsf{HS}}$	$\hat{C}^+_{\mathsf{PACE,BIC}}$	$\hat{C}^+_{\mathsf{FACE}}$	$\hat{C}^+_{\mathsf{SC}}$	$\hat{C}^+_{\mathsf{PP}}$
2	50	10	ζ_1 (BIAS)	-1.086 (0.31)	-1.600 (0.31)	-4.662 (0.27)	-1.260 (0.36)	-0.422 (0.31)	-1.328 (0.27) [90.3%]
			ζ_1 (MSE)	30.13 (2.33)	31.32 (2.43)	42.74 (2.61)	40.56 (4.79)	29.27 (2.34)	21.11 (1.54) [90.3%]
			ϕ_1 (AISE)	7.31 (0.431)	7.25 (0.402)	7.36 (0.398)	6.98 (0.457)	8.62 (0.533)	5.00 (0.375) [90.3%]
2	50	20	ζ_1 (BIAS)	-0.901 (0.30)	-1.264 (0.29)	-4.741 (0.25)	-0.605 (0.38)	-0.402 (0.30)	-0.991 (0.29) [89.3%]
			ζ_1 (MSE)	27.23 (2.02)	25.90 (1.87)	40.91 (2.52)	43.90 (9.57)	27.03 (2.04)	23.29 (1.74) [89.3%]
			ϕ_1 (AISE)	6.58 (0.780)	6.38 (0.738)	5.61 (0.366)	6.99 (0.877)	7.19 (0.813)	4.39 (0.444) [89.3%]
2	200	10	ζ_1 (BIAS)	-0.581 (0.18)	-0.680 (0.17)	-4.465 (0.15)	0.163 (0.25)	-0.400 (0.18)	-0.586 (0.15) [94.0%]
			ζ_1 (MSE)	9.65 (0.72)	9.56 (0.71)	26.85 (1.31)	19.47 (3.54)	9.46 (0.73)	6.34 (0.45) [94.0%]
			ϕ_1 (AISE)	2.12 (0.106)	2.11 (0.104)	2.79 (0.115)	2.12 (0.153)	2.19 (0.109)	1.23 (0.072) [94.0%]
2	200	20	ζ_1 (BIAS)	-0.412 (0.16)	-0.477 (0.16)	-4.473 (0.14)	0.203 (0.19)	-0.247 (0.16)	-0.236 (0.15) [90.0%]
			ζ_1 (MSE)	7.89 (0.59)	7.86 (0.59)	25.82 (1.32)	10.82 (1.27)	7.84 (0.59)	5.72 (0.42) [90.0%]
			ϕ_1 (AISE)	1.48 (0.082)	1.47 (0.080)	2.14 (0.076)	1.54 (0.123)	1.52 (0.084)	0.95 (0.067) [90.0%]
4	50	10	ζ_1 (BIAS)	-1.185 (0.31)	-1.773 (0.31)	-4.541 (0.27)	-0.871 (0.33)	-0.019 (0.32)	-1.321 (0.28) [94.0%]
			ζ_1 (MSE)	30.95 (2.82)	31.03 (2.65)	42.71 (2.50)	33.32 (3.12)	30.12 (2.97)	24.38 (2.01) [94.0%]
			ϕ_1 (AISE)	11.19 (0.799)	11.56 (0.927)	10.05 (0.538)	10.71 (0.861)	13.05 (0.952)	9.75 (0.659) [94.0%]
4	50	20	ζ_1 (BIAS)	0.036 (0.30)	-0.322 (0.29)	-4.312 (0.25)	1.521 (0.49)	0.468 (0.30)	-0.477 (0.25) [92.0%]
			ζ_1 (MSE)	26.07 (2.47)	25.61 (2.36)	37.15 (2.39)	74.96 (25.51)	27.11 (2.48)	17.85 (1.40) [92.0%]
			ϕ_1 (AISE)	7.80 (0.513)	7.40 (0.422)	6.50 (0.315)	7.70 (0.559)	8.45 (0.566)	6.04 (0.388) [92.0%]
4	200	10	ζ_1 (BIAS)	-0.173 (0.17)	-0.190 (0.17)	-4.543 (0.14)	1.042 (0.24)	-0.189 (0.17)	-0.450 (0.14) [91.3%]
			ζ_1 (MSE)	8.85 (0.69)	8.78 (0.68)	26.87 (1.35)	17.74 (2.96)	8.99 (0.71)	5.32 (0.41) [91.3%]
			ϕ_1 (AISE)	3.02 (0.116)	2.95 (0.116)	3.38 (0.128)	3.04 (0.192)	2.99 (0.116)	1.68 (0.077) [91.3%]
4	200	20	ζ_1 (BIAS)	-0.078 (0.16)	-0.093 (0.16)	-4.452 (0.13)	1.026 (0.20)	-0.074 (0.16)	-0.151 (0.14) [88.0%]
			ζ_1 (MSE)	7.37 (0.68)	7.33 (0.68)	25.04 (1.16)	12.44 (2.29)	7.29 (0.68)	5.17 (0.38) [88.0%]
			ϕ_1 (AISE)	2.06 (0.087)	1.99 (0.086)	2.48 (0.079)	2.12 (0.121)	2.05 (0.089)	1.15 (0.055) [88.0%]
10	50	10	ζ_1 (BIAS)	-1.323 (0.33)	-1.780 (0.34)	-4.739 (0.29)	-0.787 (0.39)	0.350 (0.33)	57.487 (34.76) [87.3%]
			ζ_1 (MSE)	33.91 (2.71)	37.89 (2.92)	47.38 (2.74)	47.24 (6.12)	33.50 (3.09)	>1000 [87.3%]
			ϕ_1 (AISE)	13.21 (1.172)	13.06 (1.159)	13.14 (1.130)	12.85 (1.246)	17.00 (1.372)	50.48 (4.206) [87.3%]
10	50	20	ζ_1 (BIAS)	-0.739 (0.29)	-1.064 (0.29)	-4.826 (0.25)	1.081 (0.53)	0.179 (0.30)	-1.585 (0.42) [94.0%]
			ζ_1 (MSE)	26.27 (2.25)	26.34 (2.27)	41.63 (2.52)	86.04 (23.12)	26.67 (2.48)	52.55 (6.53) [94.0%]
			ϕ_1 (AISE)	8.55 (0.465)	8.25 (0.449)	7.31 (0.360)	8.91 (0.790)	10.68 (0.540)	22.63 (2.764) [94.0%]
10	200	10	ζ_1 (BIAS)	-0.123 (0.18)	-0.173 (0.18)	-4.598 (0.16)	1.419 (0.29)	-0.065 (0.18)	-0.446 (0.24) [90.3%]
			ζ_1 (MSE)	9.70 (0.77)	9.44 (0.75)	28.56 (1.57)	26.74 (4.23)	9.59 (0.76)	16.19 (3.48) [90.3%]
			ϕ_1 (AISE)	3.45 (0.161)	3.28 (0.153)	3.59 (0.147)	5.12 (0.794)	3.58 (0.160)	6.68 (1.396) [90.3%]
10	200	20	ζ_1 (BIAS)	0.149 (0.16)	0.086 (0.16)	-4.415 (0.14)	2.810 (0.28)	0.158 (0.16)	-0.021 (0.15) [93.3%]
			ζ_1 (MSE)	7.61 (0.75)	7.60 (0.75)	24.95 (1.21)	31.26 (4.35)	7.57 (0.75)	6.11 (0.50) [93.3%]
			ϕ_1 (AISE)	2.48 (0.095)	2.26 (0.091)	2.78 (0.091)	5.65 (0.798)	2.44 (0.095)	1.78 (0.077) [93.3%]

Table S6: Similar to Table S3, but for the high error variance ($\sigma^2 = 0.1$) and $n = 50$ or 200.

L	n	m		$\hat{C}^+_{\mathsf{trace}}$	$\hat{C}^+_{\mathsf{HS}}$	$\hat{C}^+_{\mathsf{PACE,BIC}}$	$\hat{C}^+_{\mathsf{FACE}}$	$\hat{C}^+_{\mathsf{SC}}$	$\hat{C}^+_{\mathsf{PP}}$
2	50	10	ζ_2 (BIAS)	-1.8388 (0.152)	-1.7784 (0.149)	-3.5301 (0.114)	-0.9901 (0.161)	-0.7247 (0.149)	-1.2173 (0.125) [90.3%]
			ζ_2 (MSE)	10.29 (0.772)	9.80 (0.723)	16.32 (0.770)	8.77 (0.935)	7.17 (0.612)	5.73 (0.460) [90.3%]
			ϕ_2 (AISE)	12.27 (0.619)	11.33 (0.540)	14.05 (0.753)	10.57 (0.524)	13.24 (0.739)	7.37 (0.413) [90.3%]
2	50	20	ζ_2 (BIAS)	-1.3525 (0.138)	-1.3609 (0.135)	-3.5278 (0.106)	-0.7096 (0.150)	-0.6108 (0.135)	-0.8197 (0.126) [89.3%]
			ζ_2 (MSE)	7.53 (0.573)	7.29 (0.548)	15.83 (0.740)	7.27 (1.054)	5.84 (0.493)	4.89 (0.436) [89.3%]
			ϕ_2 (AISE)	8.17 (0.806)	7.86 (0.762)	7.88 (0.416)	8.60 (0.857)	9.09 (0.835)	5.52 (0.456) [89.3%]
2	200	10	ζ_2 (BIAS)	-0.7373 (0.084)	-0.6823 (0.082)	-3.2700 (0.067)	0.1459 (0.115)	-0.3201 (0.082)	-0.4211 (0.071) [94.0%]
			ζ_2 (MSE)	2.64 (0.226)	2.47 (0.213)	12.05 (0.434)	3.96 (0.761)	2.13 (0.201)	1.60 (0.134) [94.0%]
			ϕ_2 (AISE)	2.77 (0.124)	2.65 (0.119)	3.45 (0.144)	3.97 (0.273)	3.20 (0.126)	1.78 (0.081) [94.0%]
2	200	20	ζ_2 (BIAS)	-0.6300 (0.069)	-0.5852 (0.068)	-3.3402 (0.051)	-0.0371 (0.083)	-0.3411 (0.069)	-0.3683 (0.063) [90.0%]
			ζ_2 (MSE)	1.82 (0.131)	1.73 (0.127)	11.93 (0.331)	2.07 (0.262)	1.54 (0.118)	1.19 (0.095) [90.0%]
			ϕ_2 (AISE)	1.74 (0.091)	1.67 (0.088)	2.12 (0.086)	2.38 (0.124)	2.00 (0.092)	1.21 (0.070) [90.0%]
4	50	10	ζ_2 (BIAS)	-1.5477 (0.181)	-1.5070 (0.176)	-3.2394 (0.129)	-0.3727 (0.187)	0.1156 (0.171)	-0.3731 (0.144) [94.0%]
			ζ_2 (MSE)	12.22 (0.886)	11.55 (0.829)	15.47 (0.813)	10.61 (0.933)	8.71 (0.755)	5.96 (0.480) [94.0%]
			ϕ_2 (AISE)	25.22 (1.327)	22.02 (1.271)	26.39 (1.606)	18.30 (1.074)	33.37 (1.772)	30.85 (2.151) [94.0%]
4	50	20	ζ_2 (BIAS)	-0.9643 (0.144)	-1.0049 (0.144)	-3.4901 (0.104)	-0.0961 (0.157)	-0.1996 (0.138)	-0.4866 (0.123) [92.0%]
			ζ_2 (MSE)	7.15 (0.566)	7.23 (0.540)	15.42 (0.688)	7.34 (0.679)	5.76 (0.485)	4.39 (0.371) [92.0%]
			ϕ_2 (AISE)	18.89 (1.166)	15.79 (0.862)	16.58 (1.125)	14.88 (0.947)	26.21 (1.589)	19.76 (1.343) [92.0%]
4	200	10	ζ_2 (BIAS)	-0.1316 (0.092)	-0.0675 (0.090)	-3.1950 (0.071)	1.0686 (0.123)	-0.0037 (0.091)	-0.2255 (0.077) [91.3%]
			ζ_2 (MSE)	2.55 (0.203)	2.45 (0.192)	11.73 (0.456)	5.69 (0.666)	2.46 (0.196)	1.69 (0.139) [91.3%]
			ϕ_2 (AISE)	7.08 (0.428)	6.62 (0.404)	5.93 (0.307)	7.29 (0.339)	8.46 (0.551)	5.59 (0.458) [91.3%]
4	200	20	ζ_2 (BIAS)	-0.1401 (0.072)	-0.1111 (0.072)	-3.2045 (0.055)	0.9766 (0.092)	-0.0743 (0.072)	-0.1898 (0.070) [88.0%]
			ζ_2 (MSE)	1.57 (0.112)	1.55 (0.112)	11.16 (0.346)	3.48 (0.440)	1.55 (0.113)	1.31 (0.102) [88.0%]
			ϕ_2 (AISE)	4.95 (0.302)	4.60 (0.231)	3.15 (0.129)	5.71 (0.253)	5.56 (0.327)	3.33 (0.191) [88.0%]
10	50	10	ζ_2 (BIAS)	-1.5249 (0.188)	-1.3948 (0.178)	-2.9220 (0.130)	-0.2879 (0.194)	0.7590 (0.163)	0.4492 (0.921) [87.3%]
			ζ_2 (MSE)	12.93 (1.016)	11.46 (0.919)	13.61 (0.725)	11.31 (1.128)	8.51 (0.974)	221.81 (169.852) [87.3%]
			ϕ_2 (AISE)	31.04 (1.776)	27.21 (1.651)	35.65 (2.083)	22.16 (1.478)	46.32 (2.501)	79.56 (4.252) [87.3%]
10	50	20	ζ_2 (BIAS)	-0.7773 (0.164)	-0.7739 (0.158)	-3.2142 (0.116)	0.3041 (0.208)	0.4978 (0.153)	-0.4919 (0.194) [94.0%]
			ζ_2 (MSE)	8.62 (0.763)	8.04 (0.666)	14.36 (0.674)	13.03 (2.435)	7.28 (0.802)	10.87 (1.333) [94.0%]
			ϕ_2 (AISE)	22.17 (1.310)	19.24 (1.230)	20.18 (1.278)	17.50 (1.114)	33.58 (1.793)	40.76 (2.857) [94.0%]
10	200	10	ζ_2 (BIAS)	-0.1008 (0.091)	-0.0577 (0.092)	-3.0460 (0.070)	1.2327 (0.164)	0.1507 (0.090)	-0.0111 (0.113) [90.3%]
			ζ_2 (MSE)	2.50 (0.204)	2.51 (0.204)	10.76 (0.436)	9.54 (1.567)	2.45 (0.204)	3.44 (0.777) [90.3%]
			ϕ_2 (AISE)	8.58 (0.384)	7.76 (0.325)	7.84 (0.362)	10.44 (0.872)	11.30 (0.480)	13.64 (1.702) [90.3%]
10	200	20	ζ_2 (BIAS)	0.0985 (0.076)	0.0812 (0.075)	-3.0809 (0.058)	2.2418 (0.165)	0.1864 (0.074)	0.0953 (0.073) [93.3%]
			ζ_2 (MSE)	1.72 (0.152)	1.68 (0.150)	10.50 (0.339)	13.17 (1.613)	1.68 (0.152)	1.50 (0.138) [93.3%]
			ϕ_2 (AISE)	5.70 (0.284)	4.94 (0.232)	3.69 (0.161)	12.04 (1.010)	6.90 (0.349)	5.27 (0.288) [93.3%]

S3 Technical results

S3.1 Proof of Theorem 1

Proof of Theorem 1. The minimization (3) is equivalent to

$$\underset{C\in\mathcal{H}(K\otimes K)}{\arg\min}\left\{\ell(C)+\lambda\tilde{\Psi}(C)\right\},\quad \text{where}\quad \tilde{\Psi}(C)=\begin{cases}\Psi(C), & C\in\mathcal{S}^+(K)\\ \infty, & C\notin\mathcal{S}^+(K)\end{cases}.$$

For any $C\in\mathcal{H}(K\otimes K)$, let $C=C_1+C_2$ be the orthogonal decomposition in $\mathcal{H}(K\otimes K)$, where $C_1\in\mathcal{K}\otimes\mathcal{K}$ and $C_2\in(\mathcal{K}\otimes\mathcal{K})^\perp$. Note that $\ell(C)=\ell(C_1)$ as ℓ only depends on the data. Therefore, it suffices to show that $\tilde{\Psi}(C)\geq\tilde{\Psi}(C_1)$. (We define $\infty\geq\infty$.)

If $C\notin\mathcal{S}^+(K)$, $\infty=\tilde{\Psi}(C)\geq\tilde{\Psi}(C_1)$ is trivial. In the following, we assume $C\in\mathcal{S}^+(K)$. We call a $D\in\mathcal{H}(K\otimes K)$ symmetric if $D=D^\top$. We will first show that C_1 and C_2 are both symmetric, and then show that $C_1\in\mathcal{S}^+(K)$. Finally, we complete the proof by showing $\tilde{\Psi}(C)\geq\tilde{\Psi}(C_1)$.

Suppose that C is symmetric, then $C=(C_1+C_1^\top)/2+(C_2+C_2^\top)/2$. As $C_1^\top\in\mathcal{K}\otimes\mathcal{K}$ and $C_2^\top\in(\mathcal{K}\otimes\mathcal{K})^\perp$, we have $C_1=(C_1+C_1^\top)/2$ and $C_2=(C_2+C_2^\top)/2$ due to the uniqueness of orthogonal decomposition of C. Thus, C_1 and C_2 are both symmetric.

By the definition of $\mathcal{S}^+(K)$, $\mathcal{C}_C\notin\mathcal{S}^+(K)$ if there exists $f\in\mathcal{H}(K)$ such that $\langle\mathcal{C}_Cf,f\rangle_{\mathcal{H}(K)}<0$. For any $g\in\mathcal{K}$, $\langle\mathcal{C}_{C_2}g,g\rangle_{\mathcal{H}(K)}=0$, so

$$\langle\mathcal{C}_Cg,g\rangle_{\mathcal{H}(K)}=\langle\mathcal{C}_{C_1}g,g\rangle_{\mathcal{H}(K)}+\langle\mathcal{C}_{C_2}g,g\rangle_{\mathcal{H}(K)}=\langle\mathcal{C}_{C_1}g,g\rangle_{\mathcal{H}(K)}.$$

Moreover, $\langle\mathcal{C}_{C_1}h,h\rangle_{\mathcal{H}(K)}=0$ for any $h\in\mathcal{K}^\perp$. Hence $C_1\in\mathcal{S}^+(K)$ since $C\in\mathcal{S}^+(K)$.

Clearly, $\Psi(C)\geq\Psi(C_1)$ if $\tau_k(C)\geq\tau_k(C_1)$ for all k. To prove that $\tau_k(C)\geq\tau_k(C_1)$ for all k, it suffices to show $\|\mathcal{C}_Cf\|_{\mathcal{H}(K)}\geq\|\mathcal{C}_{C_1}f\|_{\mathcal{H}(K)}$ for all $f\in\mathcal{H}(K)$. Due to the fact that

$$P_{\mathcal{K}}\mathcal{C}_CP_{\mathcal{K}}=P_{\mathcal{K}}\mathcal{C}_{C_1}P_{\mathcal{K}}+P_{\mathcal{K}}\mathcal{C}_{C_2}P_{\mathcal{K}}=P_{\mathcal{K}}\mathcal{C}_{C_1}P_{\mathcal{K}}=\mathcal{C}_{C_1},$$

where $P_{\mathcal{K}}$ is the projection operator to $\mathcal{K}$, we have $\|\mathcal{C}_Cf\|_{\mathcal{H}(K)}\geq\|P_{\mathcal{K}}\mathcal{C}_CP_{\mathcal{K}}f\|_{\mathcal{H}(K)}=\|\mathcal{C}_{C_1}f\|_{\mathcal{H}(K)}$ for all $f\in\mathcal{H}(K)$. Therefore, $\tilde{\Psi}(C)\geq\tilde{\Psi}(C_1)$. □

S3.2 Proofs of Theorems 2 and 3

We first make some technical preparations before the proofs of Theorems 2 and 3. To begin with, we introduce a few notations regarding covering numbers. Following Definitions 2.2 and 2.3 in van de Geer (2000), for a class of functions $\mathcal{G}$, we denote the u-entropy of $\mathcal{G}$ for the supremum norm by $H_\infty(u, \mathcal{G})$, and the u-entropy with bracketing of $\mathcal{G}$ for $L^2(Q)$ by $H_B(u, \mathcal{G}, Q)$ where $L^2(Q) = \{g : \int |g|^2 \, dQ < \infty\}$ and Q is a probability measure. Let $\mathcal{M}$ be a metric space with the metric $\|\cdot\|_{\mathcal{M}}$. For a compact subset A of $\mathcal{M}$, we define the k-th entropy number by

$$\epsilon_k(A, \mathcal{M}) = \inf\{\epsilon > 0 : \text{there exist } g_1, \ldots, g_{2^k} \in \mathcal{M} \text{ such that } A \subseteq \cup_{j=1}^{2^k} B(g_j, \epsilon)\},$$

where $B(g_j, \epsilon) = \{g \in \mathcal{M} : \|g - g_j\|_{\mathcal{M}} \leq \epsilon\}$ represents a ball with center g_j and radius ϵ in $\mathcal{M}$.

Define $\tilde{\mathcal{F}} = \{C \in \mathcal{F} : \Psi(C) \leq 1\}$. For C such that $\Psi(C) = \sum_{k \geq 1} |\tau_k(C)|^p \leq 1$, $|\tau_k(C)| \leq 1$ for all $k \geq 1$ so $|\tau_k(C)|^2 \leq |\tau_k(C)|^p$ and

$$\|C\|^2_{\mathcal{H}(K \otimes K)} = \sum_{k \geq 1} |\tau_k(C)|^2 \leq \sum_{k \geq 1} |\tau_k(C)|^p \leq 1.$$

Therefore $\sup_{C \in \tilde{\mathcal{F}}} |C|_\infty < \infty$ due to Lemma 2.1 in Lin (2000).

Theorem 2 can be similarly established by following the exact blueprint for the proof of Theorem 3, except for the changes in Lemmas 1 and 5. The entropy in Lemma 1 would be replaced by $H_\infty(u, \tilde{\mathcal{F}}) \leq Du^{-2/r}$ by Theorem 5.2 of Birman and Solomjak (1967), and Lemma 5 would be accordingly modified by verifying a different set of conditions when Lemma 4 is applied. Therefore, hereafter we only provide the proof of Theorem 3 where $\mathcal{F} \subseteq \{C \in \mathcal{H}(K \otimes K) : C \text{ is a periodic function}\}$.

Lemma 1 (Entropy)**.** *There exists a constant $D > 0$ such that*

$$H_\infty(u, \tilde{\mathcal{F}}) \leq \left(\frac{D}{u}\right)^{1/r} \left\{\log\left(\frac{D}{u}\right)\right\}^{1+1/2r}, \quad 0 < u < D.$$

Proof. By the arguments right after the definition of the entropy number, it suffices to focus on $B_1 = \{C \in \mathcal{F} : \|C\|_{\mathcal{H}(K \otimes K)} \leq 1\}$. Due to norm equivalence and by Theorem 6.15 of Dũng et al.

(2016) ($p = 2$ in their paper), we have

$$\epsilon_k(B_1, L_\infty) \leq D\frac{(\log k)^{r+1/2}}{k^r}, \quad \text{where } D > 0 \text{ is a constant.}$$

By Lemma 4 of Cucker and Smale (2002), we have

$$H_\infty(u, \tilde{\mathcal{F}}) \leq H_\infty(u, B_1) \leq k, \quad \text{where} \quad u = D\frac{(\log k)^{r+1/2}}{k^r},$$

so

$$H_\infty(u, \tilde{\mathcal{F}}) \leq \left(\frac{D}{u}\right)^{1/r} \left\{\log\left(\frac{D}{u}\right)\right\}^{1+1/2r}, \quad 0 < u < D,$$

due to $r \geq 2$. □

Recall that in Section 5, we defined $\langle g_1, g_2\rangle_n$ and $\|g_1\|_n$ for arbitrary bivariate functions g_1 and g_2. Here we additionally define

$$\langle g_1, g_2\rangle_{n,jk} = \frac{1}{n}\sum_{i=1}^{n} g_1(T_{ij}, T_{ik})g_2(T_{ij}, T_{ik}), \quad \|g_1\|^2_{n,jk} = \langle g_1, g_1\rangle_{n,jk}, \quad 1 \leq j \neq k \leq m.$$

Note that T_{ij} and $T_{i'j}$ are not necessarily the same for $i \neq i'$. By varying $j \neq k$, we obtain $m(m-1)$ groups of n time pairs $\{(T_{ij}, T_{ik}) : i = 1, \ldots, n\}$. Below in terms of the independence between curves, we study the increment of the empirical process for each group and achieve its convergence result. We then combine the results across these groups to obtain the rate of convergence of our estimator. Note that the specific grouping does not matter in our proof, i.e., one can group different time points together as long as within-group time pairs are independent to each other.

Recall that $Z_{ijk} = Y_{ij}Y_{ik}$ as defined in Section 6.2. Additionally we define

$$\gamma_{ijk} = \gamma_i(T_{ij}, T_{ik}) = Z_{ijk} - \mathbb{E}(Z_{ijk} \mid T_{ij}, T_{ik}) = Z_{ijk} - c_{ijk}, \quad c_{ijk} = C_0(T_{ij}, T_{ik}).$$

By (9),

$$\hat{C}_\lambda = \underset{C\in\mathcal{F}}{\arg\min}\left\{\|Z - C\|^2_n + \lambda\Psi(C)\right\}. \tag{S1}$$

We begin with a basic inequality to relate the empirical norm $\|\hat{C}_\lambda - C_0\|_n$ with the empirical process $\{\sqrt{n}\langle\gamma, C - C_0\rangle_n : C \in \mathcal{F}\}$.

Lemma 2 (Basic Inequality)**.**

$$\|\hat{C}_\lambda - C_0\|_n^2 + \lambda\Psi(\hat{C}_\lambda) \le 2\langle\gamma, \hat{C}_\lambda - C_0\rangle_n + \lambda\Psi(C_0). \tag{S2}$$

Proof. By (S1), we can rewrite $\|Z - \hat{C}_\lambda\|_n^2 + \lambda\Psi(\hat{C}_\lambda) \le \|Z - C_0\|_n^2 + \lambda\Psi(C_0)$ to obtain (S2). □

The tail behavior of γ_{ijk} will be used in the subsequent proof, but it is complicated by the dependence between Y_{ij} and Y_{ik}, so we decouple the product $Y_{ij}Y_{ik}$ to obtain more manageable quantities, as shown in Lemma 3 below. A similar technique was also used in Ravikumar et al. (2011) for covariance matrix estimation.

Recall that $\mathbb{T} = \{T_{ij} : i = 1, \ldots, n; j = 1, \ldots, m\}$ and denote $\mathbb{E}_T(\cdot) = \mathbb{E}(\cdot \mid \mathbb{T}\})$.

Lemma 3 (Decoupling)**.** *Suppose that Assumptions 2–4 hold. For any pair* $(Y_{ij}, Y_{ik}), j \ne k$, *we have the decomposition*

$$Y_{ij}Y_{ik} - c_{ijk} = \frac{1}{4}(E_{ijk}^2 - e_{ijk}^2) - \frac{1}{4}(F_{ijk}^2 - f_{ijk}^2), \tag{S3}$$

where $E_{ijk} = Y_{ij} + Y_{ik}$, $F_{ijk} = Y_{ij} - Y_{ik}$, $e_{ijk}^2 = \mathbb{E}_T(E_{ijk}^2)$, *and* $f_{ijk}^2 = \mathbb{E}_T(F_{ijk}^2)$.

Moreover, conditional on $\{T_{ij} : i = 1, \ldots, n; j = 1, \ldots, m\}$, $U_{ijk} = E_{ijk}^2 - e_{ijk}^2$ *and* $V_{ijk} = F_{ijk}^2 - f_{ijk}^2$ *are both sub-exponential random variables, i.e.,*

$$2\tilde{K}_1^2\{\mathbb{E}_T \exp(|U_{ijk}|/\tilde{K}_1) - 1 - \mathbb{E}_T|U_{ijk}|/\tilde{K}_1\} \le \sigma_0^2,$$

$$2\tilde{K}_1^2\{\mathbb{E}_T \exp(|V_{ijk}|/\tilde{K}_1) - 1 - \mathbb{E}_T|V_{ijk}|/\tilde{K}_1\} \le \sigma_0^2,$$

where $\tilde{K}_1$ *and* σ_0 *are constants depending on* b_X *and* b_ε.

Proof. Obviously (S3) holds due to the fact that $\mathbb{E}_T(Y_{ij}Y_{ik}) = c_{ijk}$. Next we prove that U_{ijk} is a sub-exponential random variable. The proof for V_{ijk} is similar and is thus omitted.

By Assumption 3 and Proposition 2.1 of Rivasplata (2012), $\sup_{t\in[0,1]} \mathbb{E}\{X^2(t)\} \le b_X^2$ which implies $\sup_{s,t\in[0,1]} |C_0(s,t)| \le b_X^2$. Similarly by Assumptions 2 and 4, $\mathbb{E}_T(\varepsilon_{ij}^2) = \mathbb{E}(\varepsilon_{ij}^2) \le b_\varepsilon^2$. Then $\mathbb{E}_T(Y_{ij}^2) = \mathbb{E}_T\{X_i^2(T_{ij})\} + \mathbb{E}_T(\varepsilon_{ij}^2) \le b_X^2 + b_\varepsilon^2$ by Assumption 2. Hence $e_{ijk}^2 = \mathbb{E}_T(Y_{ij}^2) + \mathbb{E}_T(Y_{ik}^2) + 2C_0(T_{ij}, T_{ik}) \le 4b_X^2 + 2b_\varepsilon^2$.

To show that U_{ijk} possesses a sub-exponential tail, by Lemma 14.2 of Bühlmann and van de Geer (2011), it suffices to check the following moment condition: There exist positive constants K_U and σ_U such that, for all $l = 2, 3, \ldots$,

$$\mathbb{E}_T|U_{ijk}|^l \leq \frac{l!}{2} K_U^{l-2} \sigma_U^2.$$

By Proposition 3.2 of Rivasplata (2012), $\mathbb{E}_T(E_{ijk}^{2l}) \leq 2^{l+1}(l!)b^{2l}$. Due to the facts that $(x+y)^l \leq 2^l(x^l + y^l)$ and $x^l + y^l \leq 2(x+y)^l$ for $x, y > 0$,

$$\begin{aligned}\mathbb{E}_T|U_{ijk}|^l &\leq 2^l\{\mathbb{E}_T(E_{ijk}^{2l}) + e_{ijk}^{2l}\} \leq 2^l(l!)\left\{2^{l+1}b^{2l} + \frac{e_{ijk}^{2l}}{l!}\right\} \\ &\leq 2^{l+1}(l!)\left\{2^{1+1/l}b^2 + \frac{e_{ijk}^2}{(l!)^{1/l}}\right\}^l \leq 2^{l+1}(l!)\left(2^{3/2}b^2 + \frac{4b_X^2 + 2b_\varepsilon^2}{2^{1/2}}\right)^l,\end{aligned}$$

where the last inequality holds since $2^{1+1/l}$ and $1/(l!)^{1/l}$ are both decreasing in $l \geq 2$. Therefore, the moment condition above holds with properly chosen $K_U, \sigma_U^2 > 0$.

□

By Lemma 3,

$$\sup_{C\in\mathcal{F}} |\langle \gamma, C - C_0\rangle_n| \leq \frac{1}{4m(m-1)} \sum_{1\leq j\neq k\leq m} \left(\sup_{C\in\mathcal{F}} |\langle U, C - C_0\rangle_{n,jk}| + \sup_{C\in\mathcal{F}} |\langle V, C - C_0\rangle_{n,jk}| \right).$$

In view of the basic inequality (S2), it suffices to analyze $\sup_{C\in\mathcal{F}} |\langle U, C-C_0\rangle_{n,jk}|$ and $\sup_{C\in\mathcal{F}} |\langle V, C-C_0\rangle_{n,jk}|$. Our target is the increment of these empirical processes with respect to the empirical norm $\|\cdot\|_{n,jk}$. Lemma 4 below supplies a maximal inequality that will be used to obtain the increment.

Denote $Q_n = \sum_{i=1}^n \delta_{(s_i,t_i)}/n$ such that $\|g\|_{Q_n}^2 = \int g^2 dQ_n = \sum_{i=1}^n g(s_i,t_i)^2/n$.

Lemma 4. *Let $\mathcal{G}$ be a space of functions over $[0,1]^2$ and $s_i, t_i \in [0,1]$, $i = 1, \ldots, n$, be fixed time points. Suppose that $\sup_{g\in\mathcal{G}} \|g\|_{Q_n} \leq R_0$, $\sup_{g\in\mathcal{G}} |g|_\infty \leq \tilde{K}_2$, and $\{W_i : i = 1, \ldots, n\}$ are sub-exponential random variables fulfilling*

$$\max_{i=1,\ldots,n} 2\tilde{K}_1^2\{\mathbb{E}\exp(|W_i|/\tilde{K}_1) - 1 - \mathbb{E}|W_i|/\tilde{K}_1\} \leq \sigma_0^2.$$

Then if

$$\tilde{K} = 4\tilde{K}_1\tilde{K}_2,$$

$$\delta \le c_1 2R_0^2\sigma_0^2/\tilde{K}, \tag{S4}$$

$$\delta \le 8\sqrt{2}R_0\sigma_0, \tag{S5}$$

$$\sqrt{n}\delta \ge c_0\left\{\int_{\delta/2^6}^{\sqrt{2}R_0\sigma_0} H_B^{1/2}\left(\frac{u}{\sqrt{2}\sigma_0}, \mathcal{G}, Q_n\right) du \vee \sqrt{2}R_0\sigma_0\right\}, \tag{S6}$$

$$c_0^2 \ge c^2(c_1+1),$$

we have

$$\Pr\left\{\sup_{g\in\mathcal{G}}\left|\frac{1}{n}\sum_{i=1}^n W_i g(s_i,t_i)\right| \ge \delta\right\} \le c\exp\left\{-\frac{n\delta^2}{c^2(c_1+1)2R_0^2\sigma_0^2}\right\}.$$

Here c is a universal constant whereas δ, c_0, c_1 may be chosen to fulfill the above constraint.

Proof. This lemma is essentially Corollary 8.8 in van de Geer (2000), so the proof is omitted. □

With the above maximal inequality, we can apply an empirical process technique called the peeling device to obtain the increment.

Lemma 5 (Increment)**.** *Let $\mathcal{G}$ be a space of functions over $[0,1]^2$ and $s_i, t_i \in [0,1]$, $i = 1,\dots,n$, be fixed time points. Suppose that $\sup_{g\in\mathcal{G}}\|g\|_{Q_n} \le R_0$, $\sup_{g\in\mathcal{G}}|g|_\infty \le \tilde{K}_2$ and $\{W_i : i = 1,\dots,n\}$ are sub-exponential random variables fulfilling*

$$\max_{i=1,\dots,n} 2\tilde{K}_1^2\{\mathbb{E}\exp(|W_i|/\tilde{K}_1) - 1 - \mathbb{E}|W_i|/\tilde{K}_1\} \le \sigma_0^2.$$

Assume the following entropy condition holds: Let D be a constant such that $D \ge \max\{1, R_0\exp(1)\}$. For $0 < \delta \le D\exp(-1)$, we have

$$\int_0^\delta H_B^{1/2}(u,\mathcal{G},Q_n)\,du \le A_0\left\{\log\left(\frac{D}{\delta}\right)\right\}^{(1+2r)/4r}\left(\frac{\delta}{D}\right)^{1-1/2r},$$

where A_0 is a constant and $r \ge 2$.

Then, for some constants C, n_0' and T_0, depending on r, A_0, D, R_0, $\tilde{K}_1$, $\tilde{K}_2$ and σ_0, we have

$$\Pr\left\{\sup_{g\in\mathcal{G},\|g\|_{Q_n}>(\log n)^{1/2}n^{-r/(1+2r)}}\frac{|(1/n)\sum_{i=1}^n W_i g(s_i,t_i)|}{\|g\|_{Q_n}^{1-1/(2r)}} \ge T(\log n)^{(1+2r)/4r}n^{-1/2}\right\}$$
$$\le C\exp\left\{-\frac{T(\log n)^{(1+2r)/(2r)}}{C^2}\right\},$$

for all $n \geq n_0'$ *and* $T_0 \leq T \leq 4\sqrt{2}\sigma_0(\log n)^{-1/2} n^{r/(1+2r)}$.

Also, for some constants $\tilde{C}$, n_0 *and* $\tilde{T}_0$, *depending on* r, A_0, D, R_0, $\tilde{K}_1$, $\tilde{K}_2$ *and* σ_0,

$$\Pr\left\{ \sup_{g \in \mathcal{G}, \|g\|_{Q_n} \leq (\log n)^{1/2} n^{-r/(1+2r)}} \left| \frac{1}{n} \sum_{i=1}^{n} W_i g(s_i, t_i) \right| \geq \tilde{T} (\log n) n^{-2r/(1+2r)} \right\} \leq \tilde{C} \exp\left\{ - \frac{\tilde{T}(\log n)^{-1/2} n^{r/(1+2r)}}{\tilde{C}^2} \right\},$$

for all $n \geq n_0$ *and* $\tilde{T}_0 \leq \tilde{T} \leq 8\sqrt{2}\sigma_0(\log n)^{-1/2} n^{r/(1+2r)}$. *In addition,* $\tilde{T}_0 \leq T_0$ *and* $n_0 \leq n_0'$.

Proof of Lemma 5. First, we utilize Lemma 4 to develop a specialized maximal inequality as in (S11) for the rest of the proof. Let $\alpha_n = (\log n)^{(1+2r)/(4r)} n^{-1/2}$ and $(\log n)^{1/2} n^{-r/(1+2r)} \leq \omega \leq D \exp(-1)$. In Lemma 4, we replace $\mathcal{G}$ by $\mathcal{G}(\omega) = \{g \in \mathcal{G} : \|g\|_{Q_n} \leq \omega\}$ and choose $R_0 = \omega$, $\tilde{K} = 4\tilde{K}_1\tilde{K}_2$, $c_1 = A_0 \tilde{K} c_0 / (\sqrt{2}\sigma_0)$ and $\delta = 2\sigma_0^2 c_1 \omega^{1-1/(2r)} \alpha_n / \tilde{K}$. Note that, in Lemma 4, c is a universal constant. We can pick c_0 large enough such that $c_0^2 \geq c^2(c_1+1)$. We also require c_0 large enough so that $c_1 \geq 1$ and hence $c_1^2/(c_1+1) \geq c_1/2$. That is, there exists a constant $\tilde{c}_0$ such that $c_0^2 \geq c^2(c_1+1)$ and $c_1 \geq 1$ for all $c_0 \geq \tilde{c}_0$. Now, we analyze the conditions of Lemma 4:

- Condition (S4): $\delta \leq c_1 2\omega^2\sigma_0^2/\tilde{K}$. This is fulfilled due to the range constraint of ω:

$$\omega \geq \alpha_n^{2r/(1+2r)} = (\log n)^{1/2} n^{-r/(1+2r)}. \tag{S7}$$

- Condition (S5): $\delta \leq 8\sqrt{2}\omega\sigma_0$. This is satisfied if

$$c_0 \leq \frac{8}{A_0} \omega^{1/(2r)} \alpha_n^{-1} = \frac{8}{A_0} \omega^{1/(2r)} (\log n)^{-(1+2r)/(4r)} n^{1/2}. \tag{S8}$$

  Note that (S7) implies $\omega^{1/(2r)}\alpha_n^{-1} \geq \alpha_n^{-2r/(1+2r)} = (\log n)^{-1/2} n^{r/(1+2r)}$. Thus, under (S7), the requirement (S8) can be satisfied when $c_0 \leq (8/A_0)(\log n)^{-1/2} n^{r/(1+2r)}$. Clearly, for any $c_0 \geq \tilde{c}_0$, this is satisfied for sufficiently large n.

- Condition (S6):

$$\sqrt{n}\delta \geq c_0 \left\{ \int_{\delta/2^6}^{\sqrt{2}\omega\sigma_0} H_B^{1/2}\left(\frac{u}{\sqrt{2}\sigma_0}, \mathcal{G}, Q_n \right) du \vee \sqrt{2}\omega\sigma_0 \right\}.$$

  We first check $\sqrt{n}\delta \geq c_0\sqrt{2}\omega\sigma_0$, or equivalently $(\log n)^{(1+2r)/2} \geq A_0^{-2r}\omega$. This is satisfied if

$$\log n \geq \left\{ A_0^{-2r} D \exp(-1) \right\}^{2/(1+2r)}. \tag{S9}$$

Clearly, there exists a constant n_0 (independent of ω, c_0 and c_1) such that (S9) holds for all $n \geq n_0$.

We next check

$$\sqrt{n}\delta \geq c_0 \int_{\delta/2^6}^{\sqrt{2}\omega\sigma_0} H_B^{1/2}\left(\frac{u}{\sqrt{2}\sigma_0}, \mathcal{G}, Q_n\right) du,$$

which is met if

$$(\log n)^{(1+2r)/(4r)} \geq \left\{\log\left(\frac{D}{\omega}\right)\right\}^{(1+2r)/(4r)} D^{-1+1/(2r)}.$$

By (S7) and the fact $\alpha_n^{-1} < n^{1/2}$,

$$\begin{aligned}\left\{\log\left(\frac{D}{\omega}\right)\right\}^{(1+2r)/(4r)} D^{-1+1/(2r)} &\leq \left[\log\left\{\frac{D}{\alpha_n^{2r/(1+2r)}}\right\}\right]^{(1+2r)/(4r)} D^{-1+1/(2r)} \\ &= \left\{\log D + \frac{2r}{1+2r}\log(\alpha_n^{-1})\right\}^{(1+2r)/(4r)} D^{-1+1/(2r)} \\ &< \left(\log D + \frac{r}{1+2r}\log n\right)^{(1+2r)/(4r)} D^{-1+1/(2r)},\end{aligned}$$

which is a decreasing function in $D \geq 1$ if

$$\log n \geq \frac{(1+2r)^2}{2r(2r-1)}. \tag{S10}$$

We could also let n_0 such that (S10) holds for all $n \geq n_0$. Then we can ensure that

$$\left\{\log\left(\frac{D}{\omega}\right)\right\}^{(1+2r)/(4r)} D^{-1+1/(2r)} \leq \left(\frac{r}{1+2r}\log n\right)^{(1+2r)/(4r)} < (\log n)^{(1+2r)/(4r)}.$$

By Lemma 4, for all $n \geq n_0$, we have

$$\begin{aligned}\Pr\left(\sup_{g\in\mathcal{G}(\omega)}\left|\frac{1}{n}\sum_{i=1}^{n} W_i g(s_i,t_i)\right| \geq \frac{2\sigma_0^2}{\tilde{K}} c_1 \omega^{1-1/(2r)}\alpha_n\right) &\leq c\exp\left\{-\frac{2\sigma_0^2 c_1^2 \omega^{-1/r} n\alpha_n^2}{\tilde{K}^2 c^2 (c_1+1)}\right\} \\ &\leq c\exp\left(-\frac{\sigma_0^2 c_1 \omega^{-1/r} n\alpha_n^2}{\tilde{K}^2 c^2}\right),\end{aligned} \tag{S11}$$

for all $\tilde{c}_0 \leq c_0 \leq (8/A_0)(\log n)^{-1/2} n^{r/(1+2r)}$ and all $(\log n)^{1/2} n^{-r/(1+2r)} \leq \omega \leq D\exp(-1)$. This is the maximal inequality tailored for the proof of Lemma 5 and we will repeatedly use it below.

Choose $S = \min\{s \geq 1 : 2^{-s}R_0 < (\log n)^{1/2} n^{-r/(1+2r)}\}$ and $T = 2^{2-1/(2r)}\sigma_0^2 c_1/\tilde{K}$. Note that $T = 2^{3/2-1/(2r)}\sigma_0 A_0 c_0$, and therefore the condition that $\tilde{c}_0 \leq c_0 \leq (8/A_0)(\log n)^{-1/2} n^{r/(1+2r)}$ can be translated to $T_0 = 2^{3/2-1/(2r)}\sigma_0 A_0 \tilde{c}_0 \le T \le 4\sqrt{2}\sigma_0(\log n)^{-1/2} n^{r/(1+2r)}$. For $n \ge n_0$, applying the peeling device technique,

$$
\begin{aligned}
&\Pr\left\{\sup_{g\in\mathcal{G},\|g\|_{Q_n}>(\log n)^{1/2}n^{-r/(1+2r)}} \frac{|(1/n)\sum_{i=1}^n W_i g(s_i,t_i)|}{\|g\|_{Q_n}^{1-1/(2r)}} \ge T\alpha_n\right\}\\
&\le \sum_{s=1}^{S}\Pr\left\{\sup_{g\in\mathcal{G},2^{-s}R_0\le\|g\|_{Q_n}\le 2^{-s+1}R_0} \frac{|(1/n)\sum_{i=1}^n W_i g(s_i,t_i)|}{\|g\|_{Q_n}^{1-1/(2r)}} \ge T\alpha_n\right\}\\
&\le \sum_{s=1}^{S}\Pr\left\{\sup_{g\in\mathcal{G}(2^{-s+1}R_0)} \left|\frac{1}{n}\sum_{i=1}^n W_i g(s_i,t_i)\right| \ge T\alpha_n (2^{-s}R_0)^{1-1/(2r)}\right\}\\
&\le \sum_{s=1}^{S} c\exp\left\{-\frac{2^{s/r}T\alpha_n^2 n}{2^{1/(2r)+2}\tilde{K}c^2 R_0^{1/r}}\right\},
\end{aligned}
$$

where the last inequality follows from the repeated use of (S11) and holds for all $T_0 \le T \le 4\sqrt{2}\sigma_0(\log n)^{-1/2}n^{r/(1+2r)}$. If $T\alpha_n^2 n \ge 1$, the above probability is bounded by $C\exp(-T\alpha_n^2 n/C^2)$ for some constant C. Since $\alpha_n^2 n = (\log n)^{(1+2r)/(2r)}$, we can take $n \ge n_0'$ such that $T_0\alpha_n^2 n \ge 1$ and $n_0' \ge n_0$. Therefore, for all $n \ge n_0'$ and $T_0 \le T \le 4\sqrt{2}\sigma_0(\log n)^{-1/2}n^{r/(1+2r)}$, we have

$$
\Pr\left\{\sup_{g\in\mathcal{G},\|g\|_{Q_n}>(\log n)^{1/2}n^{-r/(1+2r)}} \frac{|(1/n)\sum_{i=1}^n W_i g(s_i,t_i)|}{\|g\|_{Q_n}^{1-1/(2r)}} \ge T\alpha_n\right\} \le C\exp\left\{-\frac{T(\log n)^{(1+2r)/(2r)}}{C^2}\right\}.
$$

In (S11), by choosing $\omega = (\log n)^{1/2}n^{-r/(1+2r)}$ and $\tilde{T} = 2\sigma_0^2 c_1/\tilde{K}$,

$$
\Pr\left\{\sup_{g\in\mathcal{G},\|g\|_{Q_n}\le(\log n)^{1/2}n^{-r/(1+2r)}} \left|\frac{1}{n}\sum_{i=1}^n W_i g(s_i,t_i)\right| \ge \tilde{T}(\log n)n^{-2r/(1+2r)}\right\} \le c\exp\left\{-\frac{\tilde{T}(\log n)n^{1/(1+2r)}}{2\tilde{K}c^2}\right\},
$$

for all $n \ge n_0$ and $\tilde{T}_0 := \sqrt{2}\sigma_0 A_0\tilde{c}_0 \le \tilde{T} \le 8\sqrt{2}\sigma_0(\log n)^{-1/2}n^{r/(1+2r)}$. Note that $\tilde{T}_0 \le T_0$. The range of $\tilde{T}$ covers the range of T as stated above. Also, there exists a constant $\tilde{C}$ such that the right hand side is bounded by $\tilde{C}\exp\{-\tilde{T}(\log n)^{-1/2}n^{r/(1+2r)}/\tilde{C}^2\}$. The proof is complete. □

Proof of Theorem 3. Denote $\Phi(\cdot) = \Psi^{1/p}(\cdot)$. Apparently $\Phi(\cdot)$ is a Schatten norm on $\mathcal{F}$. Let

$$
\mathcal{G} = \left\{\frac{C - C_0}{\Phi(C) + \Phi(C_0)} : C \in \mathcal{F}, \Phi(C) + \Phi(C_0) > 0\right\}.
$$

Obviously $\Psi(g) \le 1$ for all $g \in \mathcal{G}$, and $\sup_{g\in\mathcal{G}} |g|_\infty < \infty$.

By Lemma 1, for $0 < u < D$,

$$
H_\infty(u, \tilde{\mathcal{F}}) \le \left(\frac{D}{u}\right)^{1/r}\left\{\log\left(\frac{D}{u}\right)\right\}^{1+1/2r}, \quad \text{so} \quad H_\infty(u, \mathcal{G}) \le A_1\left(\frac{D}{u}\right)^{1/r}\left\{\log\left(\frac{D}{u}\right)\right\}^{1+1/2r},
$$

where A_1 is a constant. We can always choose D large enough to satisfy the entropy condition required in Lemma 5. Due to the fact that if $\log z \geq 1$,

$$\left\{-\frac{(\log z)^{(1+2r)/4r}}{z^{1-1/2r}}\right\}^{(1)} = \frac{(\log z)^{1/4r-1/2}}{z^{2-1/2r}}\left\{\left(1-\frac{1}{2r}\right)\log z - \frac{1+2r}{4r}\right\} \geq \left(\frac{1}{2}-\frac{3}{4r}\right)\frac{(\log z)^{(1+2r)/4r}}{z^{2-1/2r}},$$

we have that for δ such that $D/\delta \geq e$,

$$\int_0^\delta H_B^{1/2}(u, \mathcal{G}, \|\cdot\|_{n,jk})\, du \leq \int_0^\delta H_\infty^{1/2}(u, \mathcal{G})\, du \leq A_0 \left\{\log\left(\frac{D}{\delta}\right)\right\}^{(1+2r)/4r}\left(\frac{\delta}{D}\right)^{1-1/2r},$$

where A_0 is a constant.

Hence by Lemmas 3 and 5, for the two sets $\mathcal{G}_1 = \{g \in \mathcal{G}, \|g\|_{n,ij} \leq (\log n)^{1/2} n^{-r/(1+2r)}\}$ and $\mathcal{G}_2 = \{g \in \mathcal{G}, \|g\|_{n,ij} > (\log n)^{1/2} n^{-r/(1+2r)}\}$, we have

$$\Pr\left\{\sup_{g\in\mathcal{G}_1}\left|\frac{1}{n}\sum_{i=1}^n U_{ijk} g(s_i,t_i)\right| \geq \tilde{T}(\log n) n^{-2r/(1+2r)} \mid \mathbb{T}\right\} \leq \tilde{C}\exp\left\{-\frac{\tilde{T}(\log n)^{-1/2} n^{r/(1+2r)}}{\tilde{C}^2}\right\},$$

$$\Pr\left\{\sup_{g\in\mathcal{G}_2}\frac{|(1/n)\sum_{i=1}^n U_{ijk} g(s_i,t_i)|}{\|g\|_{n,jk}^{1-1/(2r)}} \geq T(\log n)^{(1+2r)/4r} n^{-1/2} \mid \mathbb{T}\right\} \leq C\exp\left\{-\frac{T(\log n)^{(1+2r)/(2r)}}{C^2}\right\},$$

for all $n \geq n_0'$, $T_0 \leq T \leq 4\sqrt{2}\sigma_0(\log n)^{-1/2} n^{r/(1+2r)}$, and $\tilde{T}_0 \leq \tilde{T} \leq 8\sqrt{2}\sigma_0(\log n)^{-1/2} n^{r/(1+2r)}$, where $n_0', T_0, T_1, \sigma_0, C$, and $\tilde{C}$ are all constants that do not depend on $\mathbb{T}$. This implies that both inequalities still hold when we take the supremum with respect to $\mathbb{T}$ over $\mathcal{T}^{nm}$ on each left hand side.

Therefore, we have

$$\sup_{g\in\mathcal{G}_1}|\langle U, g\rangle_{n,jk}| = \mathcal{O}_p^T\left\{\frac{\log n}{n^{2r/(1+2r)}}\right\}, \qquad \text{and} \qquad \sup_{g\in\mathcal{G}_2}\frac{|\langle U, g\rangle_{n,jk}|}{\|g\|_{n,jk}^{1-1/2r}} = \mathcal{O}_p^T\left\{\frac{(\log n)^{(1+2r)/4r}}{n^{1/2}}\right\},$$

so the following holds uniformly for all $C \in \mathcal{F}$:

$$|\langle U, C - C_0\rangle_{n,jk}| \leq \mathcal{O}_p^T\left\{\frac{\log n}{n^{2r/(1+2r)}}\right\}\{\Phi(C) + \Phi(C_0)\} + \mathcal{O}_p^T\left\{\frac{(\log n)^{(1+2r)/4r}}{n^{1/2}}\right\}\|C - C_0\|_{n,jk}^{1-1/2r}\{\Phi(C) + \Phi(C_0)\}^{1/2r}.$$

The same inequality holds for V and thus

$$|\langle U, C - C_0\rangle_{n,jk}| + |\langle V, C - C_0\rangle_{n,jk}| \leq \mathcal{O}_p^T\left\{\frac{\log n}{n^{2r/(1+2r)}}\right\}\{\Phi(C) + \Phi(C_0)\} + \mathcal{O}_p^T\left\{\frac{(\log n)^{(1+2r)/4r}}{n^{1/2}}\right\}\|C - C_0\|_{n,jk}^{1-1/2r}\{\Phi(C) + \Phi(C_0)\}^{1/2r},$$

holds uniformly for all $C \in \mathcal{F}$. Apparently

$$\frac{1}{m(m-1)} \sum_{1 \le j \ne k \le m} \|C - C_0\|_{n,jk}^{1-1/2r} \le \|C - C_0\|_n^{1-1/2r},$$

so uniformly for all $C \in \mathcal{F}$,

$$|\langle \gamma, C - C_0 \rangle_n| \le \mathcal{O}_p^T \left\{ \frac{\log n}{n^{2r/(1+2r)}} \right\} \{\Phi(C) + \Phi(C_0)\} + \mathcal{O}_p^T \left\{ \frac{(\log n)^{(1+2r)/4r}}{n^{1/2}} \right\} \|C - C_0\|_n^{1-1/2r} \{\Phi(C) + \Phi(C_0)\}^{1/2r}.$$

Therefore, the $\mathcal{O}_p^T$ result in Theorem 3 is proved following similar arguments of Theorem 10.2 of van de Geer (2000). Finally, if $S_n = \mathcal{O}_p^T(k_n)$, then $S_n = \mathcal{O}_p(k_n)$ since $\sup_{\mathbb{T} \in \mathbb{T}^{nm}} \Pr(S_n \ge Lk_n \mid \mathbb{T}) \ge \Pr(S_n \ge Lk_n)$ for all $L > 0$, and the above derivations hold if $\mathcal{O}_p^T$ is replaced by $\mathcal{O}_p$.

□

References

Birman, M. S., Solomjak, M. Z., 1967. Piecewise-polynomial approximations of functions of the classes W_p^α. Mathematics of The Ussr-sbornik 2 (3), 295.

Bühlmann, P., van de Geer, S., 2011. Statistics for high-dimensional data: methods, theory and applications. Springer, Berlin.

Cucker, F., Smale, S., 2002. On the mathematical foundations of learning. Bulletin of the American Mathematical Society 39 (1), 1–49.

Dũng, D., Temlyakov, V. N., Ullrich, T., 2016. Hyperbolic cross approximation. arXiv preprint arXiv:1601.03978.

Lin, Y., 2000. Tensor product space anova models. The Annals of Statistics 28 (3), 734–755.

Ravikumar, P., Wainwright, M. J., Raskutti, G., Yu, B., 2011. High-dimensional covariance estimation by minimizing 1-penalized log-determinant divergence. Electronic Journal of Statistics 5, 935–980.

Rivasplata, O., 2012. Subgaussian random variables: an expository note, unpublished note.

van de Geer, S., 2000. Empirical processes in M-estimation. Cambridge University Press, New York.